PCI Express® Basics

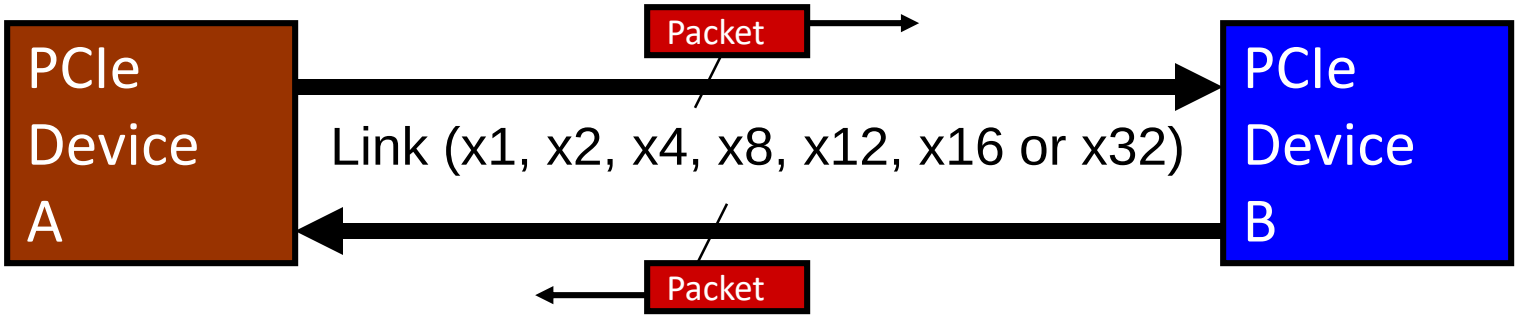
Richard Solomon
PCI-SIG® WG Member
Synopsys

Acknowledgements

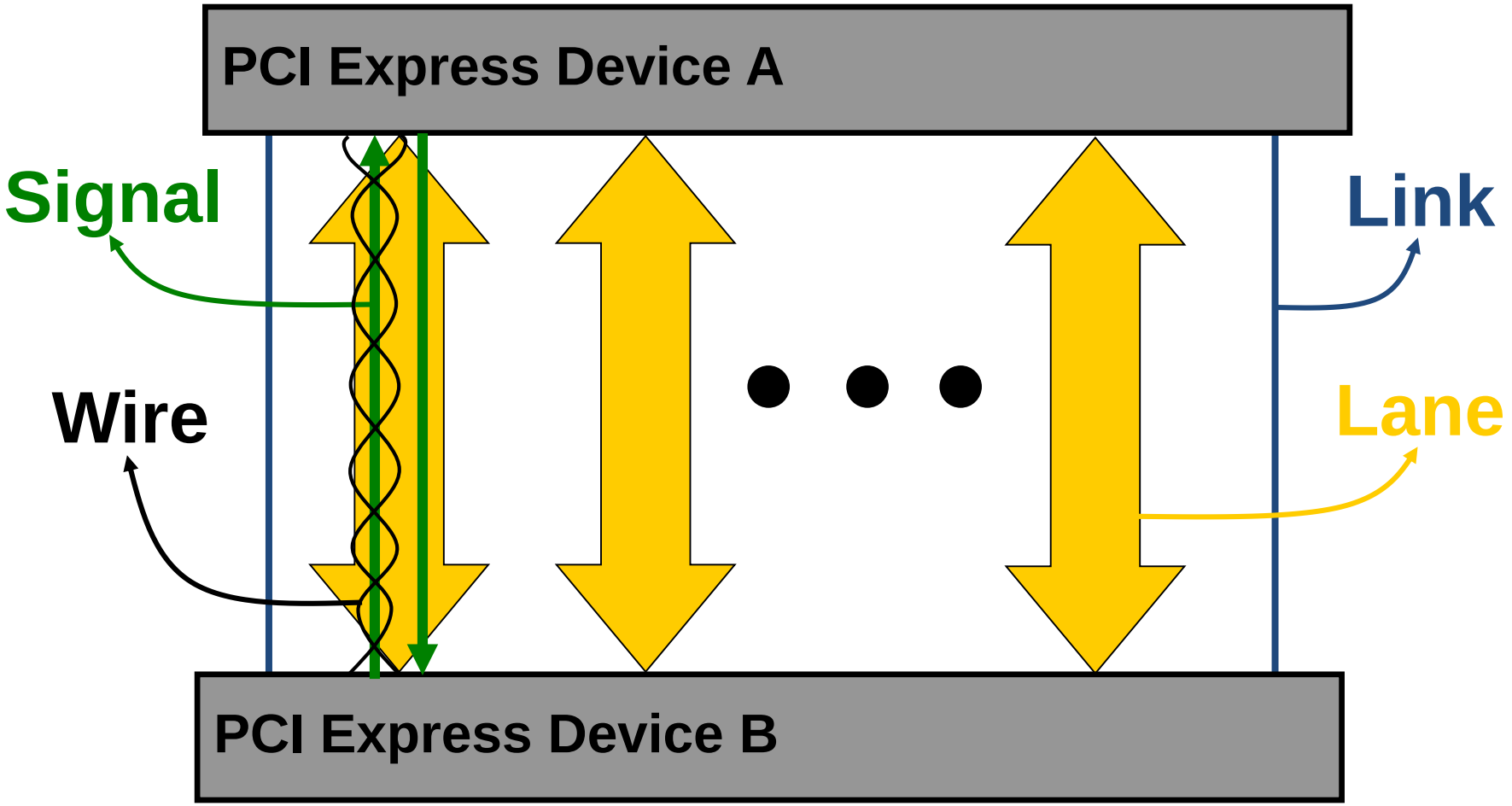
Thanks are due to Ravi Budruk, Mindshare, Inc. for much of the material on PCI Express® Basics

PCI Express Features

- **Dual Simplex point-to-point serial connection**
  - Independent transmit and receive sides
- **Scalable Link Widths**
  - x1, x2, x4, x8, *x12*, x16, *x32*
- **Scalable Link Speeds**
  - 2.5, 5.0, 8.0, 16.0 GT/s (32GT/s coming in 5.0!)
- **Packet based transaction protocol**



PCI Express Terminology



PCI Express Throughput

Bandwidth (GB/s)	Link Width				
	x1	x2	x4	x8	x16
PCIe 1.x "2.5 GT/s"	0.25	0.5	1	2	4
PCIe 2.x "5 GT/s"	0.5	1	2	4	8
PCIe 3.x "8 GT/s"	1	2	4	8	16
PCIe 4.0 "16GT/s"	2	4	8	16	32

Derivation of these numbers:

- **20% overhead due to 8b/10b encoding in 1.x and 2.x**
- **Note: ~1.5% overhead due to 128/130 encoding not reflected above in 3.x and 4.0**

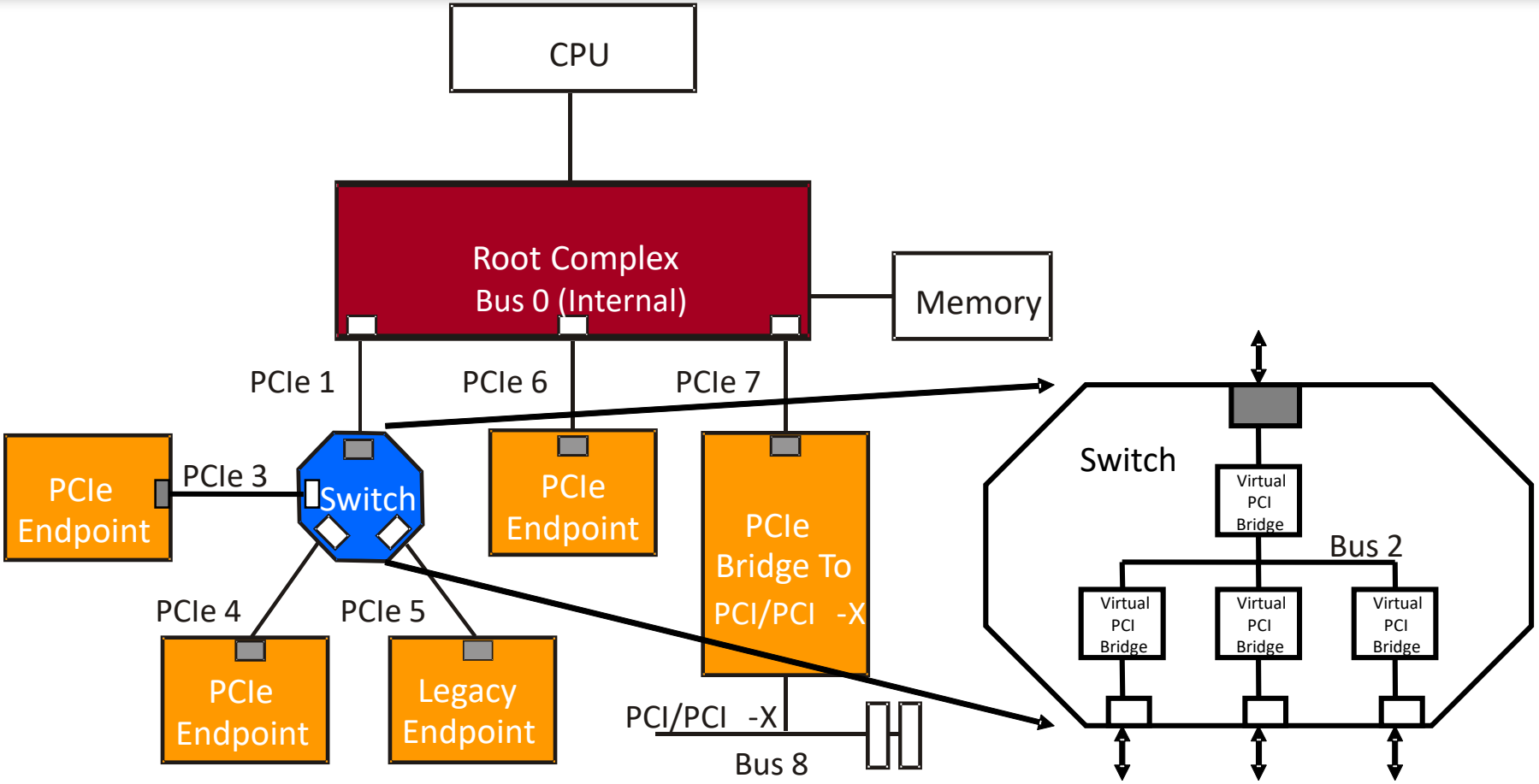
Additional Features

- **Data Integrity and Error Handling**
  - Link-level “LCRC”
  - Link-level “ACK/NAK”
  - End-to-end “ECRC”
- **Credit-based Flow Control**
  - No retry as in PCI
- **MSI/MSI-X style interrupt handling**
  - Also supports legacy PCI interrupt handling in-band
- **Advanced power management**
  - Active State PM
  - PCI compatible PM

Additional Features

- **Evolutionary PCI-compatible software model**
  - PCI configuration and enumeration software can be used to enumerate PCI Express hardware
  - PCI Express system will boot “PCI” OS
  - PCI Express supports “PCI” device drivers
  - New additional configuration address space requires OS and driver update
    - Advanced Error Reporting (AER)
    - PCI Express Link Controls

PCI Express Topology



Legend

PCI Express Device Downstream Port

PCI Express Device Upstream Port

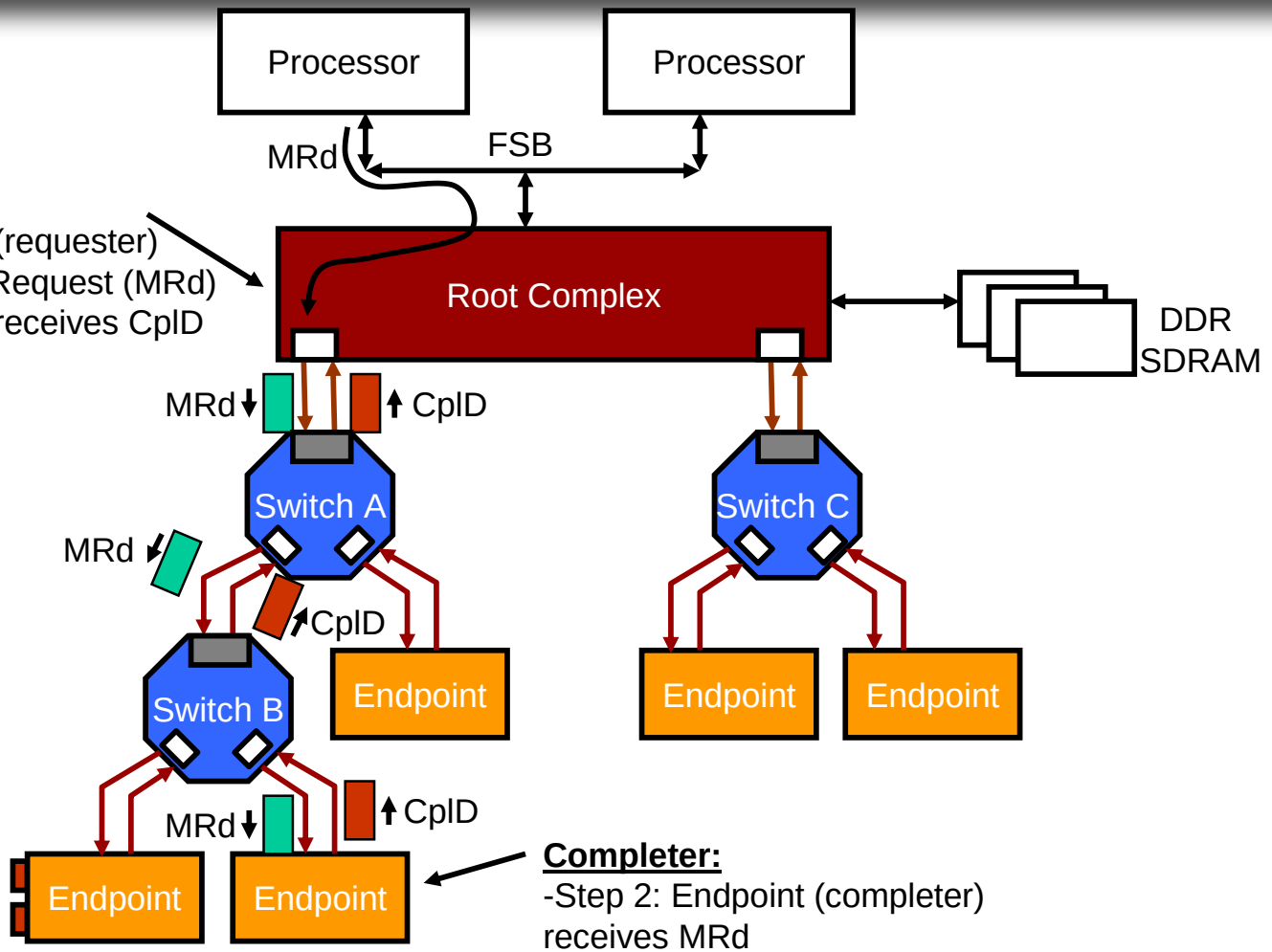
Transaction Types, Address Spaces

- **Request are translated to one of four transaction types by the Transaction Layer:**
  1. **Memory Read or Memory Write**. Used to transfer data from or to a memory mapped location.
     - The protocol also supports a *locked memory read* transaction variant
  2. **I/O Read or I/O Write.** Used to transfer data from or to an I/O location.
     - These transactions are restricted to supporting legacy endpoint devices
  3. **Configuration Read or Configuration Write**. Used to discover device capabilities, program features, and check status in the 4KB PCI Express configuration space.
  4. **Messages.** Handled like posted writes. Used for event signaling and general purpose messaging.

Three Methods For Packet Routing

- **Each request or completion header is tagged as to its *type*, and each of the packet types is routed based on one of three schemes:**
  - Address Routing
  - ID Routing
  - Implicit Routing
- **Memory and IO requests use address routing**
- **Completions and Configuration cycles use ID routing**
- **Message requests have selectable routing based on a 3-bit code in the message routing sub-field of the header type field**

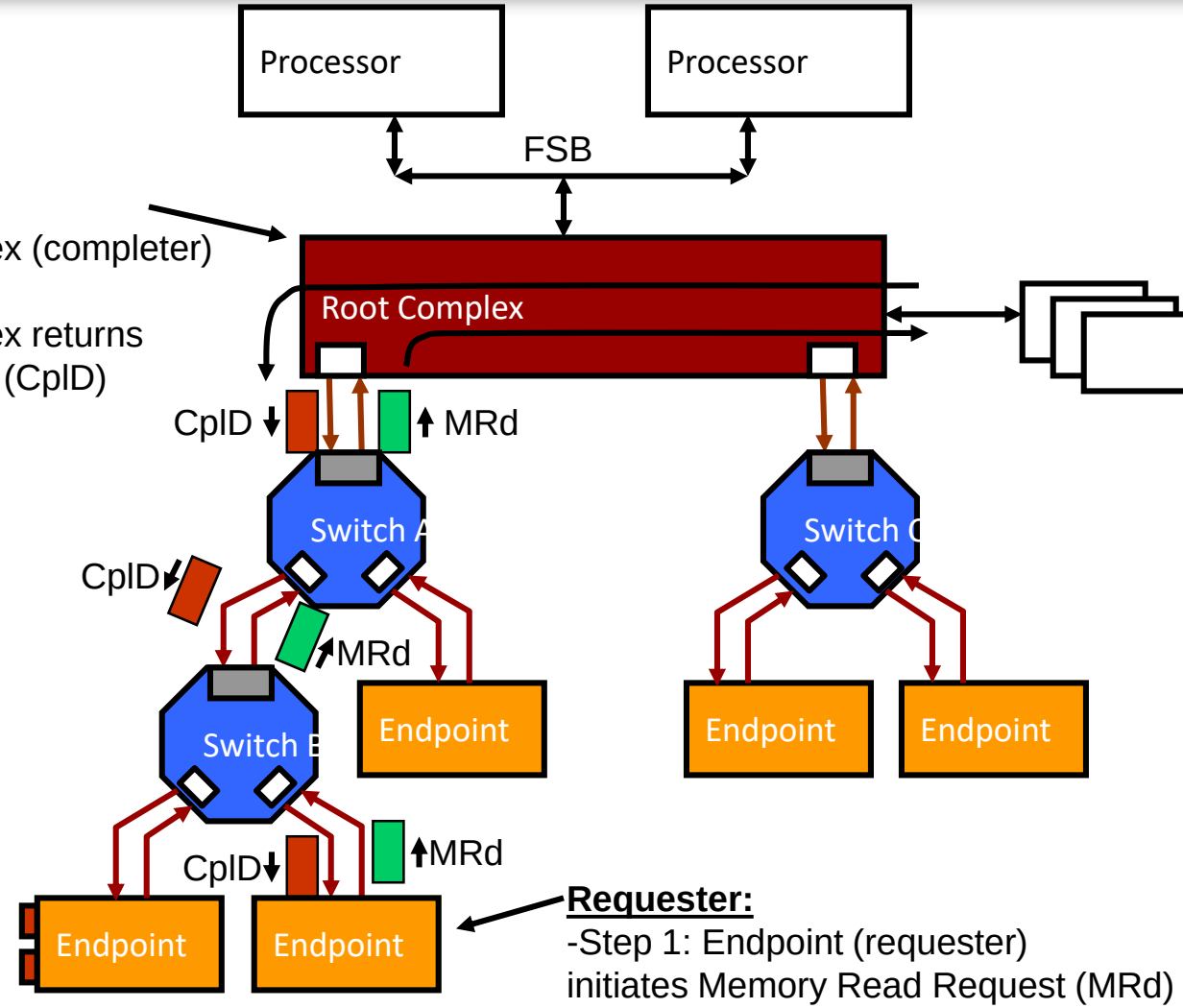
Programmed I/O Transaction



Requester:
-Step 1: Root Complex (requester) initiates Memory Read Request (MRd)
-Step 4: Root Complex receives CplD

Completer:
-Step 2: Endpoint (completer) receives MRd
-Step 3: Endpoint returns Completion with data (CplD)

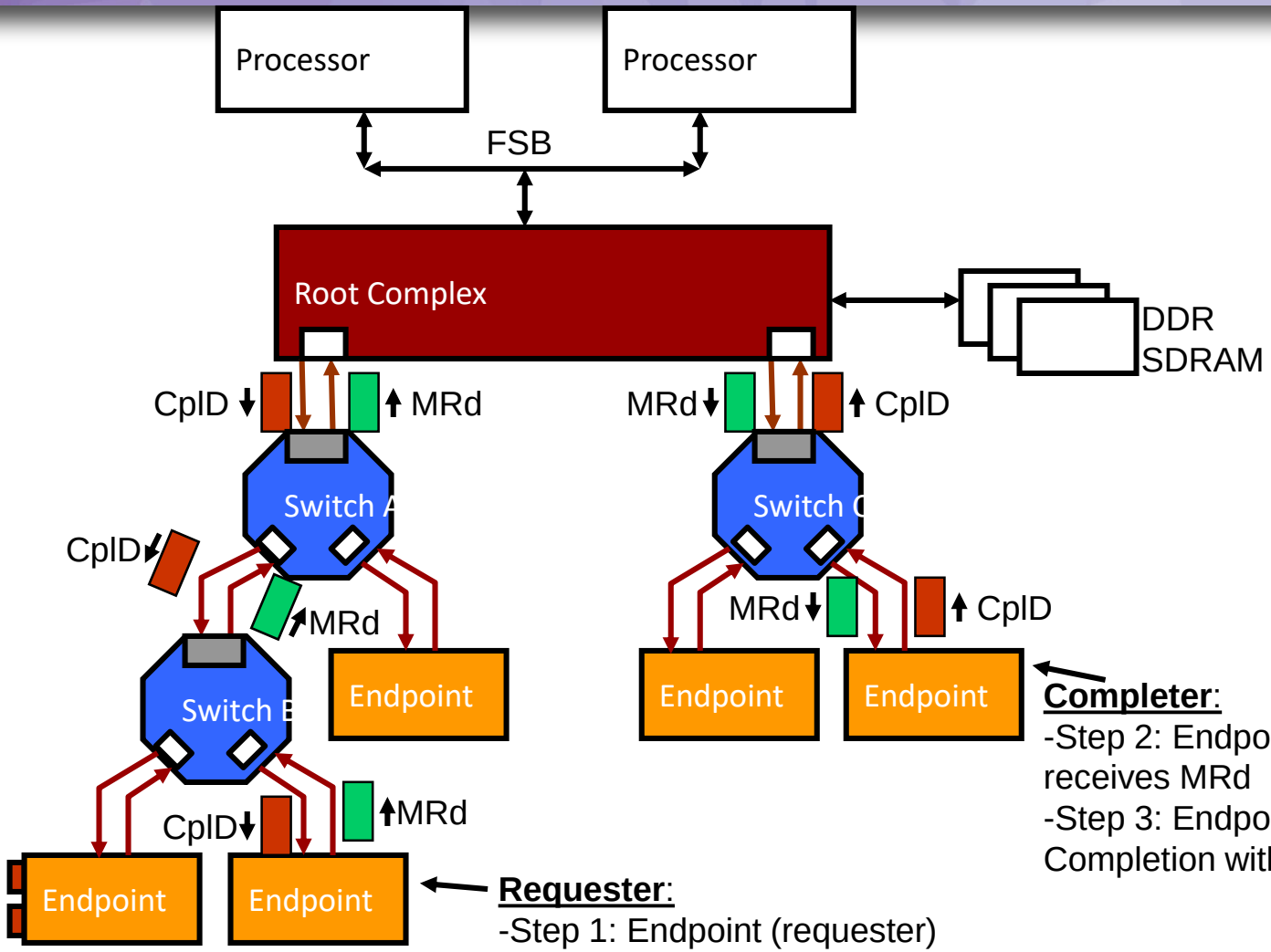
DMA Transaction



Completer:
-Step 2: Root Complex (completer) receives MRd
-Step 3: Root Complex returns Completion with data (CplD)

Requester:
-Step 1: Endpoint (requester) initiates Memory Read Request (MRd)
-Step 4: Endpoint receives CplD

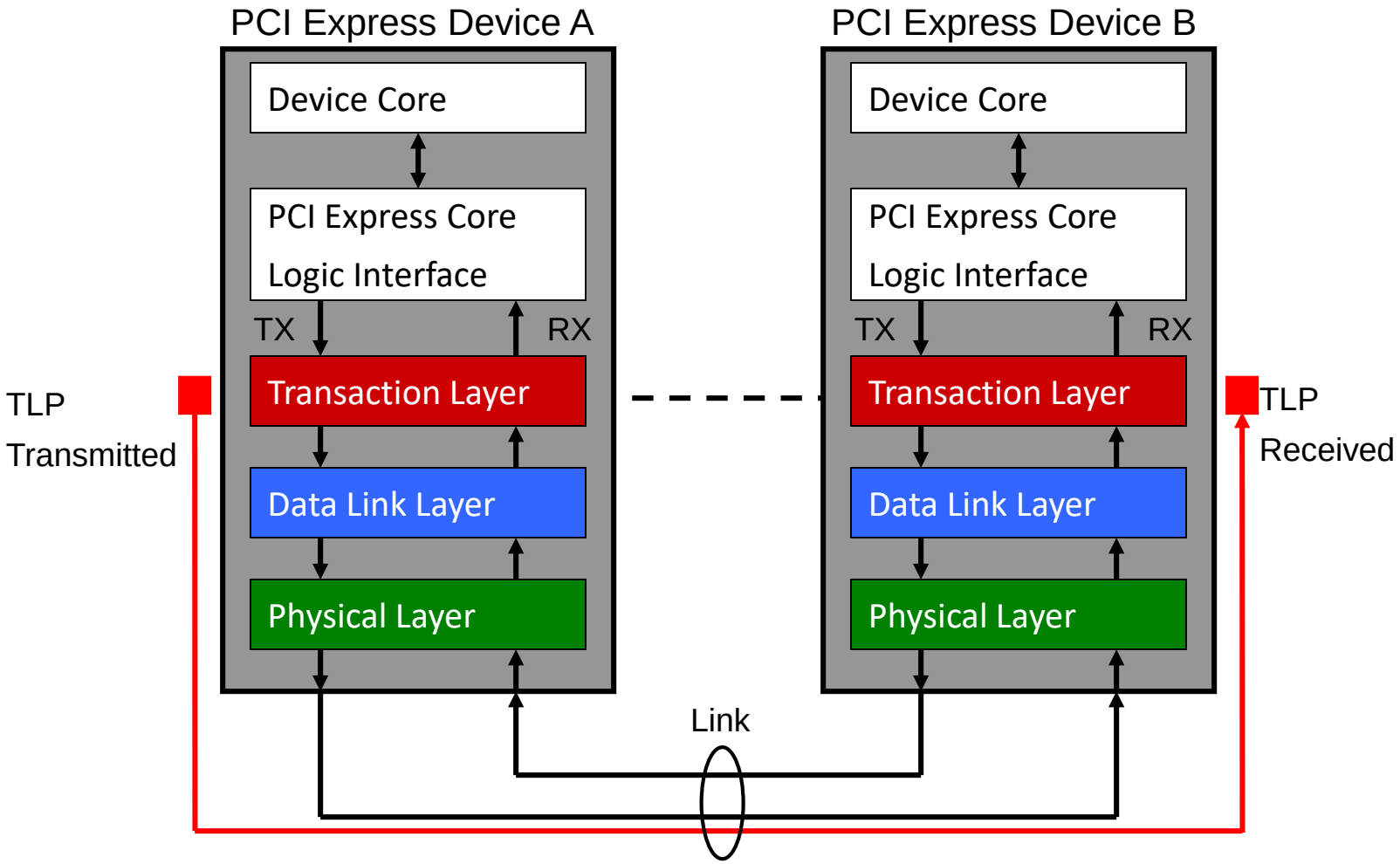
Peer-to-Peer Transaction



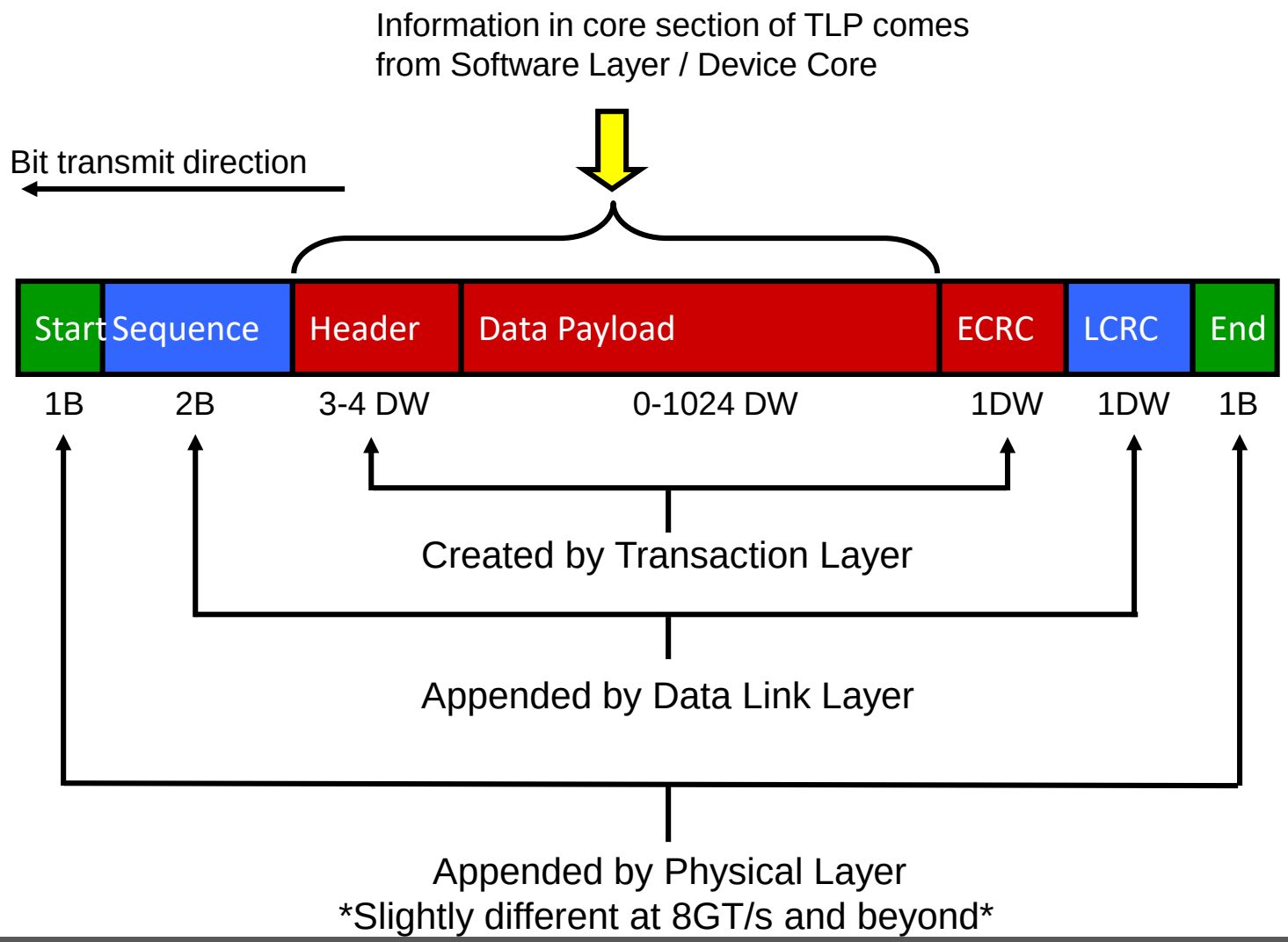
Completer:
-Step 2: Endpoint (completer) receives MRd
-Step 3: Endpoint returns Completion with data (CplD)

Requester:
-Step 1: Endpoint (requester) initiates Memory Read Request (MRd)
-Step 4: Endpoint receives CplD

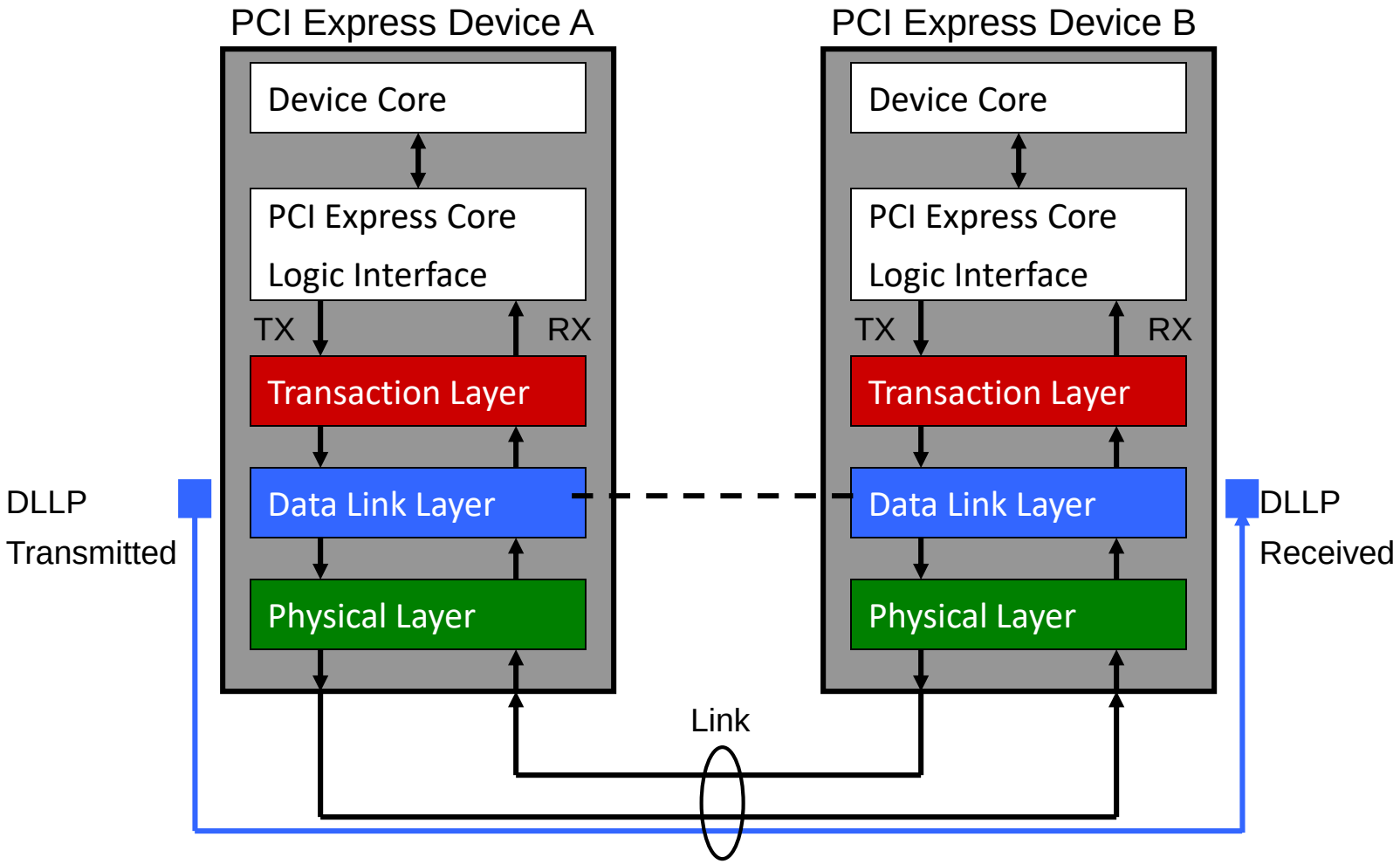
TLP Origin and Destination



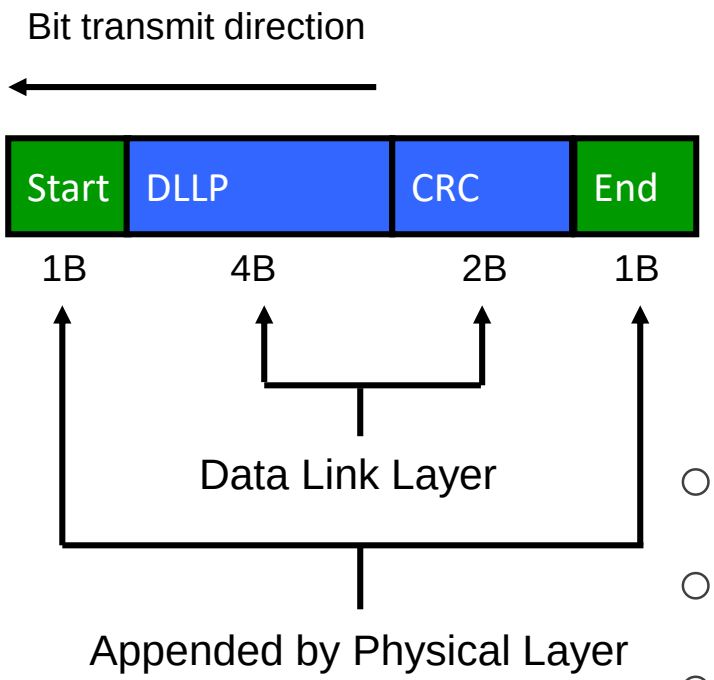
TLP Structure



DLLP Origin and Destination

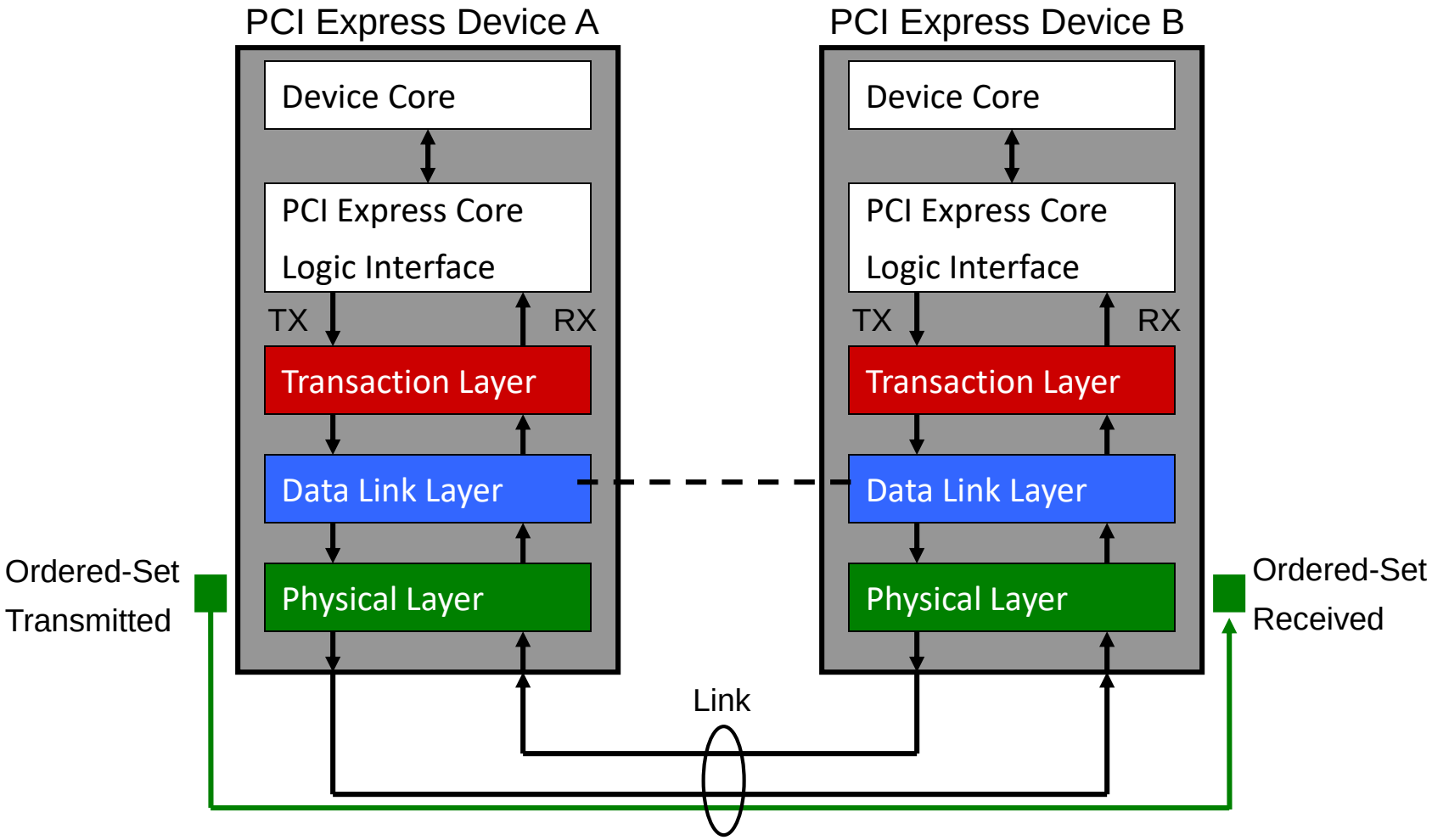


DLLP Structure



- **ACK / NAK Packets**
- **Flow Control Packets**
- **Power Management Packets**
- **Vendor Defined Packets**

Ordered-Set Origin and Destination

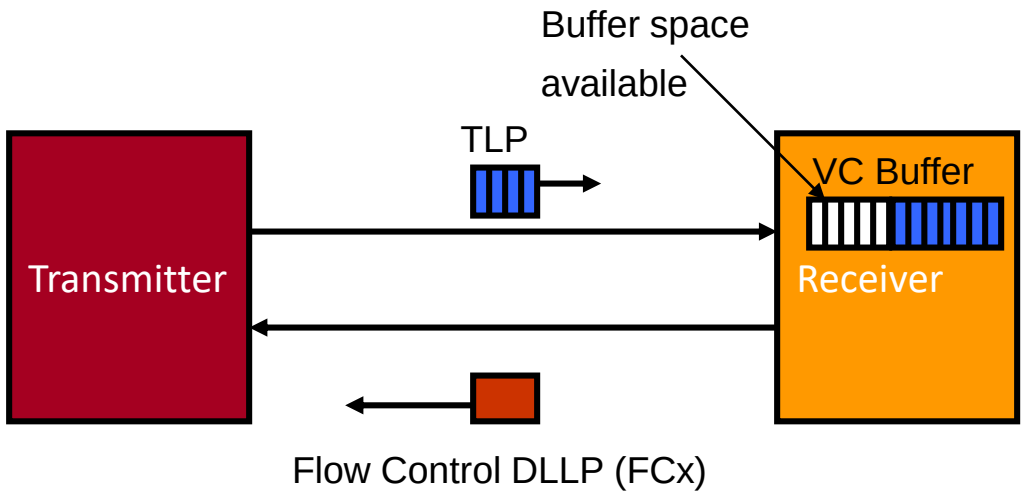


Ordered-Set Structure

- **Training Sequence One (TS1)**
  - 16 character set: 1 COM, 15 TS1 data characters
- **Training Sequence Two (TS2)**
  - 16 character set: 1 COM, 15 TS2 data characters
- **SKIP**
  - 4 character set: 1 COM followed by 3 SKP identifiers
- **Fast Training Sequence (FTS)**
  - 4 characters: 1 COM followed by 3 FTS identifiers
- **Electrical Idle (IDLE)**
  - 4 characters: 1 COM followed by 3 IDL identifiers
- **Electrical Idle Exit (EIEOS) (new to 2.0 spec)**
  - 16 characters

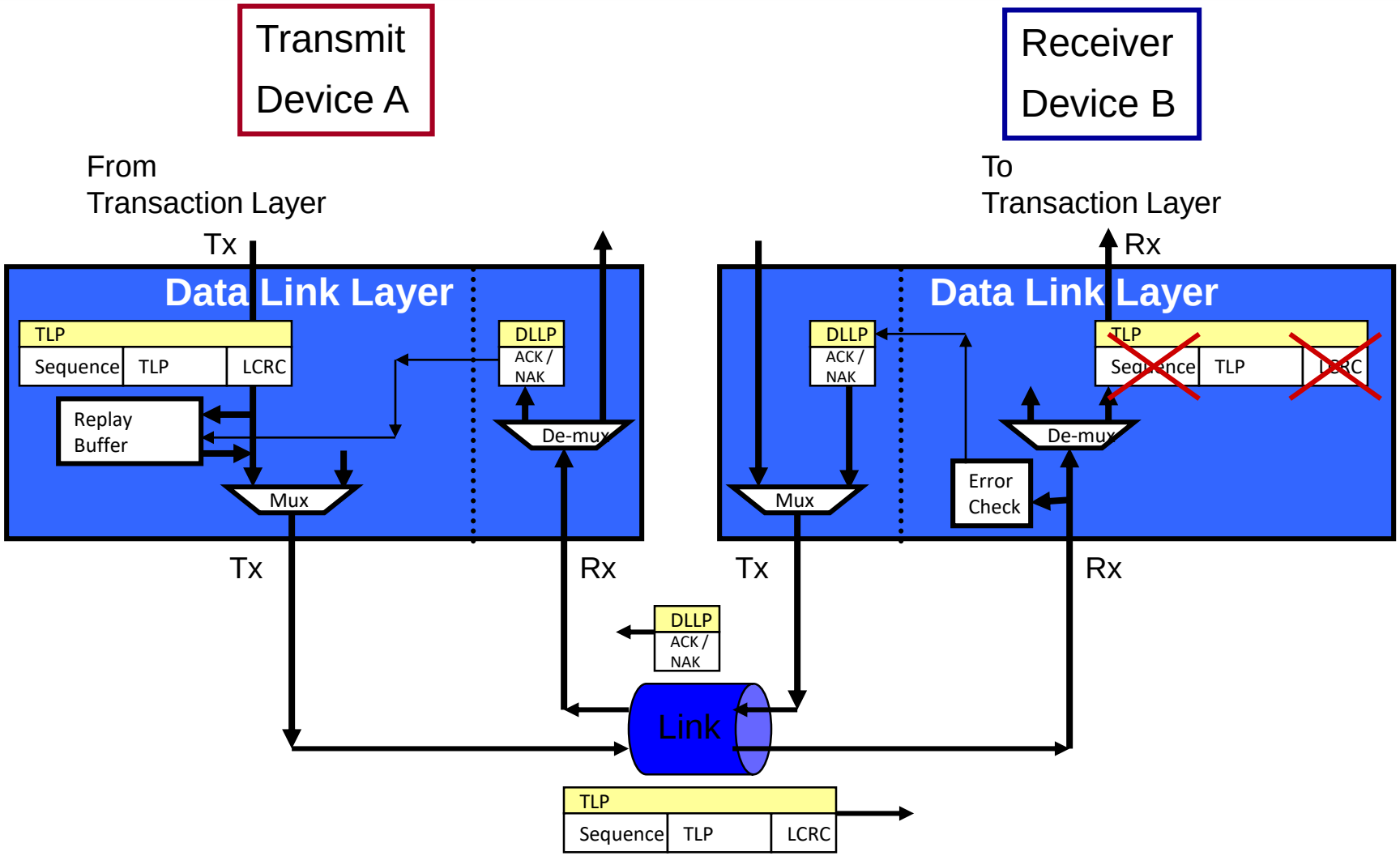
PCI Express Flow Control

Credit-based *flow control* is point-to-point based, not end-to-end



Receiver sends Flow Control Packets (FCP) which are a type of DLLP (Data Link Layer Packet) to provide the transmitter with credits so that it can transmit packets to the receiver

ACK/NAK Protocol Overview



ECRC Overview

- **"End-to-End CRC" AKA the "I Don't Trust Switches" feature**
  - Part of the TLP, therefore it's covered by the LCRC
  - Covers "invariant" parts of the TLP (almost all bits)
  - Intended for the ultimate recipient, but allowed to be checked along the way
  - Switches pass value unmodified (Multi-cast complicates)
- **Loosely defined behavior when mismatched**
  - Log and report the error like any other (including AER)
  - Requests w/bad ECRC are "*strongly recommended*" to return Unsupported Request (UR) status
  - Even credit updates are only "*strongly recommended*" on Tx/Rx of bad ECRC packet

Thank you for attending the PCI-SIG Developers Conference 2018.

For more information, please go to www.pcisig.com

Don't forget to submit your feedback via the mobile app!

Download the **Crowd Compass** app and then search for **PCI-SIG Developers Conference** or entering the following URL into your mobile browser: https://crowd.cc/s/1rKy0
Enter event code: **DevCon2018**

Alternatively, access here: https://crowd.cc/pcisig2018

Note: Create an account within the app so Admin knows who to contact if selected as the prize winner.

Each session feedback is provided is equivalent to 1 raffle entry (up to 11 sessions).
General survey feedback = 1 raffle entry.